Remote Sensing of Environment 211 (2018) 388–399

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Deriving high-quality surface emissivity spectra from atmospheric infrared sounder data using cumulative distribution function matching and principal component analysis regression

Quan Zhang[a,b], Jie Cheng[a,b,d,*], Shunlin Liang[a,b,c]

[a] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[b] Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[c] Department of Geographical Sciences, University of Maryland, College Park 20742, USA
[d] U. S. Department of Agriculture, Agricultural Research Service, Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705, USA

ARTICLE INFO

Keywords:
Emissivity
AIRS
MODIS
Cumulative distribution function
Principal component analysis

ABSTRACT

The Atmospheric Infrared Sounder (AIRS) provides limited hyperspectral thermal infrared (TIR) emissivity data for the retrieval of critical land surface and climate parameters in environmental research. However, the AIRS land surface emissivity (LSE) data lack accuracy, resulting in low-quality data retrieval, particularly for the lower boundary layer. In this study, a practical and effective method is proposed to derive high-accuracy AIRS LSE data and continuous emissivity spectra in the TIR range of 8–14.5 μm. The AIRS LSE is first rescaled to the Moderate Resolution Imaging Spectroradiometer (MODIS) LSE using cumulative distribution function (CDF) matching, and then the emissivity spectra are recovered from the rescaled AIRS LSE using principal component analysis (PCA) regression. The results show that rescaling the AIRS LSE significantly reduced the bias and root mean square (RMS) error in the study area of Africa and the Arabian Peninsula, and PCA regression successfully recovered the emissivity spectra in the 8–14.5 μm range from the rescaled AIRS LSE. At two validation sites in the Namib and Kalahari deserts of southern Africa, the biases of the rescaled AIRS LSE at three hinge points are 0.62% and 0.61%, respectively, and the biases of the recovered AIRS LSE spectra in the 8–12 μm TIR range are 0.53% and 0.56%, respectively. Variations in land cover homogeneity and the accuracy of the MODIS LSE are the critical factors impacting the final accuracy of the rescaled AIRS LSE and the recovered emissivity spectra.

1. Introduction

Land surface emissivity (LSE) is defined as the ratio of the actual emitted surface radiance to the theoretical radiance emitted from a black body of the same thermodynamic temperature (Cheng et al., 2017; Gillespie, 2014; Norman and Becker, 1995). As a measure of the amount of energy released by the land surface through radiation, LSE values normally range from approximately 0.6 to 1 in the thermal infrared (TIR) spectral range of 8–14.5 μm (Cheng and Liang, 2014a, 2014b; Prabhakara and Dalu, 1976) and are mainly influenced by the land surface composition (Xiao et al., 2003), soil moisture content (Mira et al., 2007), vegetation cover (French et al., 2008), and land relief (Mushkin and Gillespie, 2005). High-accuracy LSE data are essential to the study of the climate, land-atmosphere interactions, radiation balance, and the environment (Cheng et al., 2010; Kustas et al., 2003; Li et al., 2013; Ouyang et al., 2013).

The Atmospheric Infrared Sounder (AIRS) is a satellite-borne hyperspectral sensor. With an unprecedented 2378 sounding channels ranging from 3.75 to 15.4 μm, AIRS is an indispensable data source for short-range weather forecasting and climate research at the global scale (Aumann et al., 2003; Chahine et al., 2010; Marshall et al., 2006; Susskind et al., 2003). AIRS provides numerous geophysical parameter products, including temperature and humidity profiles, column abundances of trace gases, land and sea surface temperature and emissivity, and cloud products (Aumann et al., 2003; Susskind and Blaisdell, 2008). LSE is a critical input for retrieving land surface temperatures, atmospheric temperatures, moisture profiles, and total water vapor (Li et al., 2007; Péquignot et al., 2008; Susskind and Blaisdell, 2008; Warner et al., 2015; Yao et al., 2011). However, compared with the number of channels that are used in the retrieval of atmospheric parameters, the number of AIRS emissivity channels is relatively low. Moreover, the satellite retrieval of atmosphere parameters, such as temperature and moisture profiles in the lower boundary layer, from AIRS lacks accuracy due to the already relatively low accuracy of the

---

[*] Corresponding author at: State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China.
E-mail address: Jie_Cheng@bnu.edu.cn (J. Cheng).

https://doi.org/10.1016/j.rse.2018.04.033
Received 14 April 2017; Received in revised form 16 April 2018; Accepted 18 April 2018

AIRS LSE data (Olsen et al., 2003). Validation studies have found that the accuracy of LSE derived from AIRS is much lower than that of other sensors, such as the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and the Moderate Resolution Imaging Spectroradiometer (MODIS) (Hulley et al., 2010; Hulley et al., 2009b; Li et al., 2010; Moy et al., 2006).

An improvement in the AIRS LSE is therefore necessary because of its significance in the synergetic retrieval of atmospheric and land surface parameters. Although AIRS products have been successively improved from one version to the next through the application of modified retrieval algorithms (Susskind and Blaisdell, 2008), the accuracy of the AIRS LSE remains problematic, especially in areas with arid and semiarid climatic conditions (Hulley et al., 2010; Hulley et al., 2009b; Li et al., 2010; Moy et al., 2006). Another practical method for improving the accuracy of the AIRS LSE is blending it with a more accurate LSE dataset. Such techniques that blend the information from two sensors have already been applied to improve the accuracy of satellite-retrieved parameters (Liu et al., 2011; Peres et al., 2014; Wang and Liang, 2014). This approach was found to avoid complicated physical mechanisms while maintaining the relative data distribution pattern both temporally and spatially.

In this work, we applied cumulative distribution function (CDF) matching to improve the accuracy of the AIRS LSE by rescaling its distribution to the MODIS LSE and then used principal component analysis (PCA) regression to derive the emissivity spectra in the 8–14.5 μm TIR range from the rescaled AIRS LSE. The daytime LSE of the version 6 AIRS level 2 standard retrieval product was adopted, and the experimental period spanned 13 years from 2003 to 2015. This paper is structured as follows: the study area and datasets are described in Section 2. Data preprocessing and introduction of the proposed method are presented in Section 3. The accuracy of the rescaled AIRS LSE and regressed emissivity spectra are discussed in Section 4, and finally, conclusions are presented in Section 5.

2. Study area and data

2.1. Study area

The whole continent of Africa and the Arabian Peninsula (Fig. 1a) were selected to evaluate the proposed method, considering that these regions involve large and continuous areas of typical and homogeneous land cover types, such as desert, grassland, and forest. These land cover types are appropriate for the evaluation of coarse-resolution satellite-derived products. In addition, according to the findings of (Bannari et al., 2005; Hulley and Hook, 2009; Hulley et al., 2009a), a large pseudo-invariant desert is an ideal place for investigating the accuracy of coarse-resolution LSE products because of its spatial homogeneity. For example, the emissivity in a single 45 km pixel of the AIRS LSE data would vary only slightly in such a region, making this region an ideal validation site (Hulley et al., 2009b; Moy et al., 2006). Some recent reports (Li et al., 2012; Masiello et al., 2014; Zhang et al., 2014) documented a diurnal variation in the LSE of desert sand due to the effects of water vapor absorbed by the sand particles, which means that the LSE of desert sand is not temporally stable. The water vapor mainly affects the LSE at night; therefore, this study focused on the daytime LSE to avoid the impact of water vapor on the homogeneity of the land surface. The Namib and Kalahari deserts in southern Africa, which were used in prior studies for validating the AIRS land surface temperature and emissivity (LST&E) products (Hulley and Hook, 2009; Hulley et al., 2010; Hulley et al., 2009b), were selected as the experimental sites to validate the accuracy of the output derived from the proposed method. The Namib Desert (indicated in Fig. 1b), which is covered almost entirely with sand dunes, is adjacent to the west coast of the Atlantic Ocean and exhibits an arid climate. The Kalahari Desert (indicated in Fig. 1c), which is covered with dunes, sparse trees, and grassy shrublands, stretches across the borders of Namibia, Botswana, and South Africa and exhibits a semiarid climate. For detailed descriptions of the climate, vegetation, topography, and sand mineralogy of the Kalahari and Namib deserts, please refer to Hulley et al. (2009b).

2.2. MODIS LSE product

The MODIS/Terra LST&E product (MOD11B1) is tile-based and gridded on a sinusoidal projection. The product is provided at a spatial resolution of approximately 5 km (exactly 4.63 km) for both Version 4 (V4) and V4.1 and approximately 6 km (exactly 5.56 km) for V5 (Wan, 2008). The MOD11B1 is derived primarily using the physically based day/night algorithm, which is designed for a better estimation of LST&E, especially in arid and semiarid areas (Peres et al., 2014; Wan and Li, 1997). The change in the spatial resolution between V4 and V5 was implemented to avoid unnecessary latitudinal resampling when computing MOD11C3 climate model grids (Wan, 2008). Another change between V4 and V5 is that the emissivity for the split-window algorithm (Wan and Dozier, 1996) is partly incorporated into the V4 product but fully incorporated into the V5 product (Wan, 2008).

Hulley and Hook (2009) validated different versions of the MODIS LSE in the Namib Desert using laboratory-measured emissivity spectra from field-collected sand samples. The results indicated that V4.1 was the most accurate among V4, V4.1, and V5. The mean absolute difference was found to be 1.06% for V4, 0.65% for V4.1 and 1.93% for V5. This validation result is consistent with the findings of other independent validation studies (Cheng et al., 2014; Jacob et al., 2004). In contrast, the mean difference between the AIRS LSE and the laboratory-measured emissivity spectra reached 2.3% at the same site (Hulley et al., 2009b). The MODIS LSE product clearly performs much better than the AIRS LSE product. Therefore, the MODIS LSE was used as the reference in the CDF matching process to improve the accuracy of the AIRS LSE in this study.

Note that the MODIS LSE V4.1 data are available only since 2007, while the experimental period in this study spanned from 2003 to 2015. V4.1 is a modified but consistent version of V4. Both versions share the same retrieval algorithm, with the exception of a modified input parameter, and present similar spatiotemporal patterns, especially for arid and semiarid deserts (Hulley and Hook, 2009). Therefore, MODIS V4 was used during 2003 to 2006, and V4.1 was used thereafter.

2.3. Laboratory-measured emissivity spectra

Similar to the validation procedures of the ASTER, MODIS and AIRS LSE products in (Hulley and Hook, 2009; Hulley et al., 2009a; Hulley et al., 2009b), the laboratory-measured emissivity spectra obtained from field-collected sand samples were used to validate the rescaled AIRS LSE derived through CDF matching and the recovered emissivity spectra derived through PCA regression. In this study, the laboratory-measured emissivity spectra were carefully digitalized from the figures provided by Hulley et al. (2009b) and Hulley and Hook (2009). The procedure for obtaining the emissivity spectra from the sand samples is briefly described as follows. More than 20 sand samples from the Namib and Kalahari sites were collected during July and November 2008. The sampling locations were Sossusvlei in the Namib Desert and Tweerivieren in the Kalahari Desert (indicated in Fig. 1b and c). The spectra were then measured using a spectrometer at JPL. Given the different land cover characteristics at these two sites, the mean emissivity spectrum of pure sand samples from the Namib site was used to represent the land cover in the Namib Desert, and the weighted mean emissivity spectrum of the sand and grass samples from the Kalahari site was used to represent the land cover in the Kalahari Desert. A detailed description of the laboratory measurements can be found in Hulley et al. (2009b). Despite an inevitable change in the emissivity of the sand sample during the transfer to the laboratory due to physical disturbance (Cheng et al., 2014) and weather conditions, the derived emissivity spectra were verified to be reasonable considering the